# Topic#11
# Diagonalizability

**Recall:** Let $T \in \mathcal{L}(V)$ with $\dim(V) < \infty$.

$T$ diagonalizable $\Leftrightarrow \exists$ o.b. $\beta$ of eigenvectors of $T$

$\therefore$ diagonalizability requires existence of e-vectors

Questions: when "such" $\beta$ exist?

$1^\circ$. is there any test?

$2^\circ$. if exists, is there any way to find it out?

**Thm.** Let $T \in \mathcal{L}(V)$ with $\dim(V) = n$. Then if $T$ has $n$ distinct eigenvalues, then $T$ is diagonalizable.

**Pf.** Let $\lambda_1, \cdots, \lambda_n$ be $n$ distinct eigenvalues of $T$. For each $\lambda_i$, let $v_i$ be an eigenvector associated with $\lambda_i$. Let

$$\beta \overset{def}{=} \{v_1, \cdots, v_n\}.$$

**Claim:** $\beta$ is linearly independent. (see the pf later)

$\because \dim(V) = n = \sharp\beta$
$\therefore \beta$ is a basis for $V$. So $\beta$ is an o.b. for $V$ consisting entirely of eigenvectors of $T$. Then $T$ is diagonalizable. □

Claim is based on:

> **Lemma.** A set of eigenvectors associated with distinct eigenvalues of $T$ is linearly independent.

**Pf.:** Induction on $k \overset{def}{=} \sharp$ of such set $S$.

$k = 1$: $S = \{v_1\}, 0 \neq v_1$ is an eigenvector associated with an eigenvalue $\lambda$. Obvious to see $S = \{v_1\}$ is l. indep.

Assume "true" for $k \geq 1$, to show "true" for $k + 1$.

Let $S \overset{def}{=} \{v_1, \cdots, v_{k+1}\}$
where $v_i$ is $\lambda_i$-eVector and $\lambda_1, \cdots, \lambda_{k+1}$ distinct.

To show: $S$ l. indep.

Let $\sum_{i=1}^{k+1} a_i v_i = 0$. Apply $T - \lambda_{k+1} I$ to it, then

$$\begin{aligned} 0 &= \sum_{i=1}^{k+1} a_i (T v_i - \lambda_{k+1} v_i) \\ &= \sum_{i=1}^{k+1} a_i (\lambda_i v_i - \lambda_{k+1} v_i) \\ &= \sum_{i=1}^{k} a_i (\lambda_i - \lambda_{k+1}) v_i. \end{aligned}$$

$\because \{v_1, \cdots, v_k\}$ l. indep.
$\therefore a_1(\lambda_1 - \lambda_{k+1}) = \cdots = a_k(\lambda_k - \lambda_{k+1}) = 0$
$\because \lambda_1, \cdots, \lambda_{k+1}$ distinct
$\therefore a_1 = \cdots = a_k = 0.$
Plug to $\sum_{i=1}^{k+1} a_i v_i = 0$, then $a_{k+1} v_{k+1} = 0$
$\therefore a_{k+1} = 0$ $(v_{k+1} \neq 0)$. □□

**Warning:** The converse of Thm is false:

i.e. "if $T$ is diagonalizable then $T$ has $n$ distinct e.-Value"

NOT TRUE

**e.g.** $I_v \in \mathcal{L}(V)$ $(\dim(V) = n)$:

- diagonalizable $[I_v]_\beta = I_n$
- only one e.-value=1, $I_v(v) = 1 \cdot v$

**Let us find Necessary Conditions.**
**Observe:** Let $T \in \mathcal{L}(V)$ with $\dim(V) = n$.

$1^\circ$. $T$ has at most $n$ eigenvalues.
$2^\circ$. If $T$ is diagonanilzable, i.e. $\exists$ o.b. $\beta$ s.t.

$$[T]_\beta = D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} (\lambda_i \in \mathbb{F}),$$

then the c.p. of $T$ is given by

$$f(t) = \det(D - tI_n) = (-1)^n (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_n).$$

Thus it is necessary to require there are exactly $n$ eigenvalues counting their multiplicity!

Any other necessary conditions?

**Goal:** need compare "miltiplicity of $\lambda$" to $\dim N(T - \lambda)$!!!

**Def.** $f(t) \in P(\mathbb{F})$ **splits** over $\mathbb{F}$ **if** $\exists c$ & $a_1, \cdots, a_n$ (not necessarily distinct) in $\mathbb{F}$ such that

$$f(t) = c(t - a_1) \cdots (t - a_n).$$

e.g. if $\mathbb{F} = \mathbb{C}$, then any $f(t) \in P(\mathbb{C})$ splits over $\mathbb{C}$
e.g. if $\mathbb{F} = \mathbb{R}$, then not all $f(t) \in P(\mathbb{R})$ can split over $\mathbb{R}$, e.g. $f(t) = t^2 + 1$.

**Prop.** The c.p. of a diagonablizable $T \in \mathcal{L}(V)$ over $\mathbb{F}$ must split over $\mathbb{F}$.

**Pf.** See the previous observation. □

**Observe:** If the c.p. $f(t)$ splits, i.e.

$$f(t) = c(t - a_1) \cdots (t - a_n),$$

then we may also rewrite it as:

$$f(t) = c(t - a_1)^{m_1}(t - a_2)^{m_2} \cdots (t - a_k)^{m_k}$$
$$a_1, a_2, \cdots, a_k\text{: distinct in } \mathbb{F} \ (k \leq n)$$
$$m_1, m_2, \cdots, m_k \geq 1 : m_1 + \cdots + m_k = n$$

**Def.:** Let $\lambda \in \mathbb{F}$ be an eigenvalue of $T \in \mathcal{L}(V)$ with c.p. $f(t)$. Then, the **algebraic multiplicity** of $\lambda$ is defined to be the largest positive integer $k$ for which $(t - \lambda)^k$ is a factor of $f(t)$.

e.g. Let $m_\lambda$ denote the a.m. of $\lambda$, then $m_{a_i} = m_i$.

Consider the following issue: If c.p.

$$f(t) = c(t-\lambda_1)^{m_1}(t-\lambda_2)^{m_2}\cdots(t-\lambda_k)^{m_k},$$

$\lambda_1, \cdots, \lambda_k$: distinct eigenvalues, $m_i =$ a.m. of $\lambda_i$, $1 \le i \le k$,

then can we know anything on

$$N(T - \lambda_i I_V)$$

in particular, on its dim (**geometric multiplicity of $\lambda_i$**)?

**We will show:**

$1^\circ$. $1 \le dim N(T - \lambda I_v) \le m_\lambda$

$2^\circ$. (i) $f_T(t)$ splits, (ii) $dim N(T - \lambda_i I_v) = m_{\lambda_i}$, $1 \le i \le k$

If (i) and (ii) both hold, then T is diagonalizable.

**Def.** Let $\lambda$ be an eigenvalue of $T \in \mathcal{L}(V)$.

$$E_\lambda \stackrel{def}{=} \{v \in V : T(V = \lambda v)\} = N(T - \lambda I_V),$$

is called the **eigenspace** of $T$ associated with $\lambda \in \mathbb{F}$.

**Lemma.** $1 \leq \dim(E_\lambda) \leq m_\lambda$.

**Proof.** Note that $E_\lambda$ is a subspace of $V$ containing at least one nonzero vector (an eigenvector associated with $\lambda \in \mathbb{F}$), then

$$1 \leq \dim(E_\lambda) \leq \dim(V) \stackrel{def.}{=} n.$$

Let $p \stackrel{def}{=} \dim(E_\lambda)$, and $\{v_1, \cdots, v_p\}$ be an o.b. for $E_\lambda$.
Extend $\{v_1, \cdots, v_p\}$ to o.b. $\beta = \{v_1, \cdots, v_p, v_{p+1}, \cdots, v_n\}$ for $V$.

Note: For $i = 1, \cdots, p$,

$$0 \neq v_i \in E_\lambda = N(T - \lambda I), \text{ i.e., } T(v_i) = \lambda v_i.$$

$$\therefore A \overset{def.}{=} [T]_\beta = \begin{pmatrix} \lambda I_p & \vdots & B \\ \cdots & \cdots & \cdots \\ 0 & \vdots & C \end{pmatrix}_{n\times n} \qquad \text{for some B and C}$$

(Get directly from
$[T]_\beta = ([T(v_i)]_\beta | \cdots | [T(v_p)]_\beta | [T(v_{p+1})]_\beta | \cdots | [T(v_n)]_\beta))$

$$\therefore \text{ c.p. of } T : f(t) = \det(A - tI_n) = \det \begin{pmatrix} (\lambda - t)I_p & \vdots & B \\ \cdots & \cdot & \cdots \\ 0 & \vdots & C - tI_{n-p} \end{pmatrix}$$

$$= \det((\lambda - t)I_p) \cdot \det(C - tI_{n-p})$$

$$= (\lambda - t)^p \cdot g(t), \quad \text{for some } g \in P_{n-p}(\mathbb{F})$$

$\therefore \dim(E_\lambda) = p \leq m_\lambda =$ algebraic multiplicity of $\lambda$. □

**The next goal:** Let $T \in \mathcal{L}(V)$, $\dim(V) = n$ with c.p.

$$f(t) = (-1)^n (t - \lambda_1)^{m_1} \cdots (t - \lambda_k)^{m_k}$$

where $\lambda_1, \cdots, \lambda_k$: distinct, and $m_1 + \cdots + m_k = n$.
We know:

$$1 \leq \dim(E_{\lambda_i}) \leq m_i,\ i = 1, \cdots, k.$$

**to show the Thm on the next page:**

**Thm.** Let $T \in \mathcal{L}(V)$ with $\dim(V) < \infty$. Assume that the c.p. of T splits over $\mathbb{F}$ and $\lambda_1, \cdots, \lambda_k$ are all the distinct eigenvalues of $T$. Then,

(a) $T$ is diagonalizable **iff**

$$m_{\lambda_i} = \dim(E_{\lambda_i}) \text{ for all } i = 1, \cdots, k;$$

(b) If $T$ is diagonalizable and $\beta_i$ is an o.b. for $E_{\lambda_i}$ $(1 \leq i \leq k)$, then

$$\beta \overset{def}{=} \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k \text{ is an o.b. for } V \text{ consisting of e-vectors of } T.$$

An example of (b) of the Thm will be presented later (the Example.3)

**Lemma:** Let $T \in \mathcal{L}(V)$ with $\dim(V) < \infty$,
$\lambda_1, \cdots, \lambda_k$ be distinct eigenvalues of $T$,
$S_1, \cdots, S_k$ be (finite) l. indep. subsets of $E_{\lambda_1}, \cdots, E_{\lambda_k}$, resp.
Then,

$$S \stackrel{def}{=} S_1 \cup \cdots \cup S_k \subset V \quad \text{is l. indep.}$$

**Pf of Lemma:** Set $n_i = \sharp S_i$ and $S_i = \{v_{i1}, \cdots, v_{in_i}\} \subset E_{\lambda_i}$.
Then, $S = \cup_{i=1}^k S_i = \{v_{ij} : 1 \le i \le k, 1 \le j \le n_i\}$.
To show $S$ is l. indep., let $\sum_{i=1}^k \sum_{j=1}^{n_i} a_{ij} v_{ij} = 0$,
rewrite it as $0 = \sum_{i=1}^k w_i$, where each $w_i \overset{def}{=} \sum_{j=1}^{n_i} a_{ij} v_{ij} \in E_{\lambda_i}$.

**Claim:** $w_1 = \cdots = w_k = 0$.

If claim is true, then $0 = \sum_{j=1}^{n_i} a_{ij} v_{ij}$ $(1 \le i \le k)$.
Note, $S_i$ is l. indep. for each $i$,
hence all $a_{ij} = 0$ $(1 \le i \le k, 1 \le j \le n_i)$. Thus S is l.indep. □

**Pf of Claim:** Otherwise, some $w_i$ is nonzero.
Remove those zero vectors in $\sum_{i=1}^k w_i$, and renumber $w_i$, we have

$$w_1 + \cdots + w_m = 0 \text{ (each } w_i \in E_{\lambda_i} \text{ is nonzero)},$$

For $1 \le i \le m$, by definition, $w_i$ is an e-vector of $\lambda_i$.
So, this is a contradiction to "a set of eigenvectors of distinct e-values must be l. indep." □□

**Pf of the Thm.** Let $n = \dim(V)$, $m_i = m_{\lambda_i}$, $d_i = \dim(E_{\lambda_i})$, $1 \le i \le k$.

**Pf of (a):** "$\Rightarrow$" Assume: $T$ is diagonalizable.
$V$ has an o.b. $\beta$ of e-vectors of $T$, set $\beta_i = \beta \cap E_{\lambda_i}, 1 \le i \le k$.
We see $\sharp\beta_i \le d_i \le m_i$ $(1 \le i \le k)$, then

$$n = \sharp\beta = \sum_{i=1}^{k} \sharp\beta_i \le \sum_{i=1}^{k} d_i \le \sum_{i=1}^{k} m_i = n.$$

The second equality is from 'disjoint' of $\beta_i$
$\therefore \sum_{i=1}^{k}(m_i - d_i) = 0$ (note: $m_i - d_i \ge 0$ for each $i$)
$\therefore$ $m_i = d_i$, $1 \le i \le k$.

"$\Leftarrow$" Assume: $m_i = d_i$ $(1 \le i \le k)$.
Let $\beta_i$ be an o.b. for $E_{\lambda_i}$, set $\beta = \beta_1 \cup \cdots \cup \beta_k$.
Note: $\beta$ is l. indep. and

$$\sharp\beta = \sum_{i=1}^{k} \sharp\beta_i = \sum_{i=1}^{k} \dim(E_{\lambda_i}) = \sum_{i=1}^{k} d_i = \sum_{i=1}^{k} m_i = n.$$

$\because$ $dim(V) = n$ $\therefore$ $\beta$ is an o.b. for $V$ consisting of eigenvectors of $T$.
$\therefore$ $T$ is diagonablizable.

**Pf of (b)**: direct consequence of the proof of "$\Leftarrow$" in (a). □

**Sum. Test for Diagonablization:**

Let $T \in \mathcal{L}(V)$ with $\dim(V) = n$.

Then, $T$ is diagonalizable **iff**
1°. The c.p. of $T$ splits
2°. For each eigenvalue $\lambda$ of $T$,

$$\underbrace{m_\lambda}_{\text{algebraic multiplicity of } \lambda} = \underbrace{\dim(E_\lambda)}_{\text{geometric multiplicity of } \lambda}$$

Note: $\dim(E_\lambda) = n - \text{rank}(E_\lambda)$.

Remark: For 2°, if $m_\lambda = 1$, then 2° always holds true, because in this case

$$1 \leq \dim(E_\lambda) \leq m_\lambda = 1, \text{ then } m_\lambda = \dim(E_\lambda).$$

**Example 1.** Let $A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix} \in M_{3\times 3}(\mathbb{R})$.
Determine its diagonalizability.

$1^\circ$. $f_A(t) = \det(A - tI_3) = \det \begin{pmatrix} 3-t & 1 & 0 \\ 0 & 3-t & 0 \\ 0 & 0 & 4-t \end{pmatrix}$
$= \cdots = -(t-4)(t-3)^2.$
$\therefore$ The c.p. $f_A(t)$ of $A$ splits.

$2^\circ$. $\lambda_1 = 4, m_{\lambda_1} = 1, \therefore$ 2nd condition is satisfied for $\lambda_1$.
$\lambda_2 = 3, m_{\lambda_2} = 2.$

$A - \lambda_1 I_3 = A - 3I_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ with rank$= 2$.
$\therefore \dim(E_{\lambda_2}) = 3 - 2 = 1 < 2 = m_{\lambda_2}$
$\therefore$ 2nd condition fails for $\lambda_2$.
Therefore A is NOT diagonalizable. □

**Example 2.** Let $T : P_2(\mathbb{R}) = P_2(\mathbb{R})$,

$$f \mapsto T(f), T(f(x)) = f(1) + f'(0)x + (f'(0) + f''(0))x^2.$$

(1) Note $T \in \mathcal{L}(P_2(\mathbb{R}))$.
Let $\alpha = \{1, x, x^2\}$: s.o.b. Compute

$T(1) = 1$
$T(x) = 1 + 1 \cdot x + (1 + 0)x^2 = 1 + x + x^2$
$T(x^2) = 1 + 0 \cdot x + (0 + 2)x^2 = 1 + 2x^2$

$$\therefore [T]_\alpha = \left(\begin{array}{c|c|c} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{array}\right).$$

□

(2) Test diagonalization for $T$:

Let $f_T(t) = \det([T] - tI_3) = \det \begin{pmatrix} 1-t & 1 & 1 \\ 0 & 1-t & 0 \\ 0 & 1 & 2-t \end{pmatrix}$

$= \cdots = -(t-1)^2(t-2)^1.$

$\therefore f_T(t)$ splits.

$\lambda_1 = 1 : m_{\lambda_1} = 2$, $[T]_\alpha - \lambda_1 I = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$ with rank$= 1$.

$\therefore \dim(E_{\lambda_1}) = 3 - 1 = 2 = m_{\lambda_1}$.

$\lambda_2 = 2$, $m_{\lambda_2} = 1 = \dim(E_{\lambda_2})$.

Therefore $T$ is digonablizable. □

(3) Goal: Find an o.b. $\beta$ of $P_2(\mathbb{R})$ consisting of e-vectors of $T$

$$\text{so that } [T]_\beta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Idea: $T(v) = \lambda v \, (v \neq 0) \Leftrightarrow [T]_\alpha [v]_\alpha = \lambda [v]_\alpha \, ([v]_\alpha \neq 0)$.
Thus, goal above is equivalent to find an o.b. $\gamma$ of $\mathbb{R}^3$ consisting of e-vectors of $[T]_\alpha$, then

$$\beta \overset{def}{=} \Phi_\alpha^{-1}(\gamma).$$

Specifically,
$\lambda_1 = 1$ :

$$E_{\lambda_1} = N([T]_\alpha - 1 \cdot I_3) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0 \right\}$$

$$\therefore \gamma_1 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\} \text{ is a basis for } E_{\lambda_1}.$$

$\lambda_2 = 2:$

$$E_{\lambda_2} = N([T]_\alpha - 2 \cdot I_3) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : \begin{pmatrix} -1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0 \right\}$$

$\therefore \gamma_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$ is a basis for $E_{\lambda_2}$.

Let

$$\gamma = \gamma_1 \cup \gamma_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

$\therefore \gamma$ is an o.b. for $\mathbb{R}^3$ consisting of eigenvectors of $[T]_\alpha$. □

Set

$$\beta = \Phi_\alpha^{-1}(\gamma) = \{1, -x + x^2, 1 + x^2\},$$

which is an o.b. for $P_2(\mathbb{R})$ consisting of eigenvectors of $T$. □□

**Sum.**

$$\begin{array}{ccc} V & \xrightarrow{\quad T \quad} & V \\ \Big\downarrow \Phi_\alpha & & \Big\downarrow \Phi_\alpha \\ \mathbb{F}^n & \xrightarrow{\quad [T]_\alpha \quad} & \mathbb{F}^n \end{array}$$

1°. Given a diagonable $T \in \mathcal{L}(V)$, find a convenient o.b. $\alpha$ for $V$ and work on $[T]_\alpha$, i.e. find an o.b. $\gamma$ of $\mathbb{F}^n$ consisting of eigenvectors of $[T]_\alpha$.

2°. Define

$$\beta \stackrel{def.}{=} \Phi_\alpha^{-1}(\gamma),$$

then $\beta$ is an o.b. for $V$ of eigenvectors of $T$ ($\because \Phi_\alpha : V \to \mathbb{F}^n$ is an isomorphism), so that $[T]_\beta$ is a diagonal matrix with diagonal entries given by the corresponding e-values.

**Example 3:** Let $A \in M_{n\times n}(\mathbb{F})$. Assume that $A$ is diagonalizable. Then, $f_A(t)$ splits. Let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be distinct e-values. Let $\gamma_1, \cdots, \gamma_k$ be o.b.'s for e-spaces $E_{\lambda_1}, \cdots, E_{\lambda_k}$, resp. Note

$$m_{\lambda_i} = \dim(E_{\lambda_i}), \quad n = \sum_{i=1}^{k} m_{\lambda_i}.$$

$\gamma \overset{def.}{=} \gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$: o.b. for $\mathbb{F}^n$ of eigenvectors of A.

$$\therefore [L_A]_\gamma = \begin{pmatrix} \overbrace{\lambda_1 \; \ddots \; \lambda_1}^{m_1 \text{ terms}} & & \\ & \ddots & \\ & & \overbrace{\lambda_k \; \ddots \; \lambda_k}^{m_k \text{ terms}} \end{pmatrix} \overset{def}{=} D.$$

On the other hand, from Topic#9 (page 6),

$$Q \overset{def}{=} (\underbrace{\square \cdots \square}_{\gamma_1} \mid \underbrace{\square \cdots \square}_{\gamma_2} \mid \cdots \mid \underbrace{\square \cdots \square}_{\gamma_k}) \in M_{n\times n}(\mathbb{F}).$$

$$(\sharp\gamma_k = \dim(E_{\lambda_k}) = m_k, \quad \sum \sharp\gamma_k = n)$$

$[L_A]_\gamma = Q^{-1}AQ$. ($Q = [I]_\gamma^{s.o.b.}$ changing $\gamma$-coor. to s.o.b. coor.)

$$\therefore\ Q^{-1}AQ = D, \text{ i.e. } A = QDQ^{-1}.$$

It is then easier to compute $A^n\,(n = 1, 2, \cdots)$ as

$$A^n = QD^nQ^{-1}.$$

(Only need to compute ${\lambda_i}^n$ for $1 \le i \le k$)